

ay



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/007,615	11/07/2001	Jeffrey S. Kopal	BOC9-2001-0039 (284)	4195
40987	7590	11/15/2005	EXAMINER	
AKERMAN SENTERFITT P. O. BOX 3188 WEST PALM BEACH, FL 33402-3188			ALBERTALLI, BRIAN LOUIS	
			ART UNIT	PAPER NUMBER
			2655	
DATE MAILED: 11/15/2005				

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

10/007,615

Applicant(s)

KOBAL ET AL.

Examiner

Brian L. Albertalli

Art Unit

2655

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 29 August 2005.
2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-23 is/are pending in the application.
4a) Of the above claim(s) _____ is/are withdrawn from consideration.
5) ☐ Claim(s) _____ is/are allowed.
6) ☒ Claim(s) 1-23 is/are rejected.
7) ☐ Claim(s) _____ is/are objected to.
8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Continued Examination Under 37 CFR 1.114

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on August 29, 2005 has been entered.

Response to Amendment

2. The amendments to the claims have been entered. Claims 1-3, 5, 6, 10, 15-17, 19, and 20 are currently amended.

Response to Arguments

3. Applicant's arguments filed August 29, 2005 have been fully considered but they are not persuasive.

Baker et al. disclose a method/system for generating pronunciations. As pointed out by the Applicant, a large portion of Baker et al. describes two possible methods for the generation of pronunciations (see page 12, 1st and 2nd paragraph of Applicant's arguments). Both methods involve the user entering a spelling of a new word that will be added to the dictionary, then pronouncing that word so that a speech recognizer can produce an initial estimate of the phonetic pronunciation of the new word. The applicant

has alleged that these two methods disclosed by Baker et al. "obviate not only the need, but the opportunity, to compose the pronunciation from individual phonemes corresponding to activatable visual identifiers" (see page 12, 1st paragraph, lines 7-10 and 2nd paragraph, lines 6-10 of Applicant's arguments).

The Examiner disagrees with this assessment of Baker et al. because, though brief, Baker et al. disclose a third method of generating a pronunciation wherein the user can optionally **directly type the pronunciation into box 1756** (of Fig. 17, see column 18, lines 5-6). Directly typing the pronunciation (as opposed to using one of the two methods described above) is equivalent to the Applicant's claimed "composing" as it is the addition, phoneme-by-phoneme, of phonemes that will make up the pronunciation of the new word. As admitted by the Applicant, Baker et al. further disclose editing a pronunciation presented in pronunciation box 1756, which as described in previous rejections, comprises selectively adding and removing individual phonemes to further refine the pronunciation.

Having been established, then, that Baker et al. disclose "composing" pronunciations in an equivalent manner as the present invention, there is the remaining question of whether the composing is done through "activatable visual identifiers corresponding to individual ones of a plurality of phonemes". The Applicant has described a "word history" button 1770 that allows a user to add and delete words (see page 13, 3rd paragraph of Applicant's arguments). It is noted that the "word history" button 1770 has not been proposed by the Examiner at any point to meet the "activatable visual identifier" limitation. Rather, Baker et al. disclose a phoneme table

button 68 that opens a table containing valid phonemes (visual identifiers corresponding to individual ones of a plurality of phonemes, column 18, lines 49-51). Baker et al. further disclose that the pronunciation box can be edited (which includes directly typing into the pronunciation box, and thus “composing”) using the phoneme table (making the phoneme table an “activatable” visual identifier, column 18, lines 51-52).

Therefore, for the above reasons, the Examiner maintains that Baker et al. anticipate the composition of a pronunciation of a portion of text through the selection of activatable visual identifiers corresponding to the individual ones of a plurality of phonemes.

Further, regarding the Applicant’s arguments that Baker et al. do not permit composing a pronunciation at least partially based upon an audible rendering of a portion of a portion of the pronunciation during the user’s composing the pronunciation without compiling the pronunciation information, as described above, Baker et al. disclose a user-directed composing of the pronunciation with individual phonemes. Further, the text-to-speech button 1762 plays back whatever is in pronunciation box 1756, regardless of the state of composing or editing (column 18, lines 43-45). That is, pressing the text-to-speech button will provide a pronunciation of whatever is in pronunciation box 1756, including a portion of a pronunciation (see Fig. 17, the pronunciation in box 1756 is missing the initial “j” phoneme that should be produced by the first “G” in “George”). Therefore, the text-to-speech button gives the user the capability to compose a pronunciation “based upon at least one of an audible rendering

of a portion of said pronunciation". This is equivalent to the present invention's process for playing back whatever pronunciation is in window 210 (see page 9, lines 13-14 of specification).

4. Therefore, for the reasons given above, the rejections made in the previous Office Action stand.

Claim Rejections - 35 USC § 102

5. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

6. Claims 1-4, 9-10, 12-13, 15-18, and 23 are rejected under 35 U.S.C. 102(b) as being anticipated by Baker et al. (U.S. Patent 6,092,044).

In regard to claims 1 and 15, Baker et al. discloses a computer implemented method for composing a pronunciation of a portion of text (column 17, lines 66-67 and column 18, lines 5-8) by generating pronunciation information and machine-readable storage, having stored thereon a computer program having a plurality of code sections executable by a machine for causing the machine to perform the steps of:

graphically presenting at least one activatable visual identifier corresponding to individual ones of a plurality of phonemes (Fig. 17, control window 1750 includes a

phoneme table button 68 that opens a table containing valid phonemes, column 18, lines 49-51);

responsive to a selection of one of said visual identifiers, generating said pronunciation information in accordance with said selected visual identifier (pronunciation box 1756 is edited using the phoneme table, column 18, lines 51-52), said pronunciation information comprising at least one of a phoneme selected from said plurality of phonemes (phonemes "o", "UH", "th", etc. in box 1756, selected from the phoneme table containing valid phonemes, column 18, lines 49-52), an ordering of selected phonemes (the phonemes in box 1756 are presented in the order that they are pronounced), a pronunciation stress parameter, and a prosodic parameter;

enabling a user to compose said pronunciation by selectively performing at least one of adding a particular one of the plurality of phonemes and removing a particular one of the plurality of phonemes (phonemes contained in pronunciation box 1756, column 18, lines 5-6 and lines 51-52; editing encompasses the inserting, removing, and reordering of information), said user's selection being based upon said pronunciation information and based upon at least one of an audible rendering of a portion of said pronunciation during said user's composing said pronunciation and without compiling said pronunciation information (see explanation above in Response to Arguments), audible rendering of an exemplary word illustrative of a particular phoneme, and a visual rendering of an exemplary word illustrative of the particular phoneme (audibly playing back the phonemes through a text-to-speech synthesizer, column 18, lines 43-45);

compiling said pronunciation information responsive to a selection of one of said plurality of visual identifiers (the pronunciation is added to the dictionary, column 18, lines 7-10).

In regard to claims 2-4 and 16-18, Baker et al. disclose that the phoneme table is used to edit the pronunciation information (phonemes contained in pronunciation box 1756, column 18, lines 51-52). Editing encompasses the inserting, removing, and reordering of information. The phoneme associated with the selected visual identifier would necessarily be the phoneme the user intended to insert or remove from the phoneme information. Therefore, Baker et al. disclose the step of identifying at least one phoneme associated with said selected visual identifier and inserting said identified at least one phoneme into said pronunciation information, the step of identifying at least one phoneme associated with said selected visual identifier and removing said identified at least one phoneme into said pronunciation information, as well as the step of reordering a plurality of phonemes of said pronunciation information.

Further, Baker et al. disclose the inserting and removing is based on an audible rendering of a portion of said pronunciation during said user's composing said pronunciation and without compiling said pronunciation information (see explanation above in Response to Arguments).

In regard to claim 9 and 23, Baker et al. disclose storing the pronunciation information in memory (add button 1758 adds words to the vocabulary, column 18, lines 7-10).

In regard to claim 10, Baker et al. disclose a pronunciation composition tool comprising:

A library comprising a plurality of phonemes (dictionary, column 18, lines 7-10);

A graphical user interface comprising a plurality of activatable visual identifiers corresponding to particular ones of said plurality of phonemes (phoneme table, column 18, lines 49-51); and

A processor configured to generate pronunciation information by including selected ones of said plurality of phonemes from said library responsive to a selection of at least one of said activatable visual identifiers (phoneme table contains valid phonemes, column 18, lines 49-51, used to edit pronunciation box 1756, column 18, lines 51-52, phonemes in pronunciation box generated by a processor column 19, lines 5-10) and by enabling a user to compose said pronunciation by selectively causing said processor to perform at least one operation of adding a particular one of the plurality of phonemes and removing a particular one of the plurality of phonemes (phonemes contained in pronunciation box 1756, column 18, lines 51-52; editing encompasses the inserting, removing, and reordering of information) said user causing said processor to perform at least one operation based upon said pronunciation information and at least one of an audible rendering of a portion of said pronunciation during said user's

Art Unit: 2655

composing said pronunciation and without compiling said pronunciation information (see explanation above in Response to Arguments), an audible rendering of an exemplary word illustrative of a particular phoneme, and a visual rendering of an exemplary word illustrative of a particular phoneme.

In regard to claim 12, Baker et al. disclose a compiler (processor, column 19, line 7) that compiles the pronunciation information for use with a speech driven application (once completed, the pronunciation is added to the dictionary of the speech recognizer, column 18, lines 7-10).

In regard to claim 13, the processor is further configured to modify the pronunciation information (the user can edit pronunciations in pronunciation box 1756, column 18, lines 5-6).

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

8. Claims 5-6, 8, 14, 19-20, and 22 are rejected under 35 U.S.C. 103(a) as being unpatentable over Baker et al., in view of Shaw et al. (U.S. Patent 6,363,342).

Art Unit: 2655

In regard to claims 5 and 19, Baker et al. do not disclose changing at least one parameter of said pronunciation information.

Shaw et al. disclose a method of generating pronunciation information that comprises a graphically presented means for pronunciation information by changing a pronunciation stress parameter (Fig. 2, stress buttons 50 alter the stress applied to the syllable, column 4, lines 32-36).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Baker et al. to include a parameter in the pronunciation information and to change the parameter of pronunciation information, as disclosed by Shaw et al., so the word represented by the pronunciation information would be pronounced correctly in a text-to-speech converter, thereby increasing the intelligibility of the audibly output word. Additionally, if the pronunciation information were to be used to generate models for a speech recognition device, changing the parameter of the pronunciation information would conform the recognition models more closely to input speech, thereby increasing recognition results.

In regard to claim 6 and 20, the combination of Baker et al. and Shaw et al., as applied to claim 5, above, discloses in Shaw et al. that the parameter consists of a stress parameter and a prosodic parameter (stress is a prosodic parameter, i.e. prosody refers to the intonation, rhythm, and vocal stress of speech, column 4, lines 32-36).

In regard to claim 8, 14, and 22, Baker et al. does not disclose the plurality of phonemes includes phonemes from at least two languages.

Shaw et al. discloses a plurality of phonemes includes phonemes form at least two languages (phonetic dictionaries contain phonemes corresponding to a plurality of languages, column 4, lines 11-25).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Baker et al. to include phonemes from at least two languages in order to facilitate the development of word pronunciations in the users native language, as taught by Shaw et al. (column 4, lines 23-25).

9. Claims 7, 11, and 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Baker et al., in view of Holm et al. (U.S. Patent 5,850,629).

In regard to claims 7 and 21, Baker et al. discloses playing an audio approximation of said pronunciation information (text-to-speech button plays back phonemes in pronunciation box 1756).

Baker et al. does not disclose playing an audio approximation of said pronunciation information responsive to a selection of one of said plurality of visual identifiers.

Holm et al. discloses a method of generating pronunciation information that comprises a graphically presented means for cycling through available phonemes and playing an audio approximation of those phonemes (column 7, line 66 through column 8, line 8).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Baker et al. to play an audio approximation of the pronunciation information in response to a selection of a visual identifier of that pronunciation information so that a user who was not familiar with phonetic representations could hear the sound produced by the selected phoneme, as taught by Holm et al. (column 7, line 66 through column 8, line 6).

In regard to claim 11, Baker et al. discloses a text-to-speech system configured to play an audio approximation of said pronunciation information (column 18, lines 43-45).

Baker et al. does not disclose the text-to-speech system is configured to play an audio approximation of said pronunciation information responsive to activation of one of said activatable visual identifiers.

Holm et al. discloses a text-to-speech system (Fig. 1, 36) configured to cycle through available phonemes and playing an audio approximation of those phonemes (column 7, line 66 through column 8, line 8).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the text-to-speech system of Baker et al. to play an audio approximation of the pronunciation information in response to a selection of a visual identifier of that pronunciation information so that a user who was not familiar with phonetic representations could hear the sound produced by the selected phoneme, as taught by Holm et al. (column 7, line 66 through column 8, line 6).


Conclusion

10. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Dick (U.S. Patent 4,831,654) discloses an interface for editing pronunciation parameters in a text-to-speech dictionary.

11. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brian L. Albertalli whose telephone number is (571) 272-7616. The examiner can normally be reached on Mon - Fri, 8:00 AM - 5:30 PM, every second Fri off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Wayne Young can be reached on (571) 272-7582. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).


W. R. YOUNG
PRIMARY EXAMINER